2023-09-25

A Deep Learning-Based Object Detection Framework for Automatic Asphalt Pavement Patch Detection Using Laser Profiling Images

Ibrahim Hassan Syed
Technological University Dublin, ibrahim.h.syed@mytudublin.ie

Susan McKeever Dr.
Technological University Dublin, susan.mckeever@tudublin.ie

Kieran Feighan
Pavement Management Services Ltd, Ireland, kfeighan@pms.ie

See next page for additional authors

Recommended Citation

Syed, I.H., McKeever, S., Feighan, K., Power, D., O'Sullivan, D. (2023). A Deep Learning-Based Object Detection Framework for Automatic Asphalt Pavement Patch Detection Using Laser Profiling Images. In: Christensen, H.I., Corke, P., Detry, R., Weibel, JB., Vincze, M. (eds) Computer Vision Systems. ICVS 2023. Lecture Notes in Computer Science, vol 14253. Springer, Cham. DOI: 10.1007/978-3-031-44137-0_18

Funder: Science Foundation Ireland

Authors
Ibrahim Hassan Syed, Susan McKeever Dr., Kieran Feighan, David Power, and Dympna O'Sullivan

A deep learning-based object detection framework for automatic asphalt pavement patch detection using laser profiling images.

Syed Ibrahim Hassan [1[0000-0002-0480-989X]], Susan McKeever[1[0000-0003-1766-2441]], Kieran Feighan[2], David Power[2] and Dympna O'Sullivan[1[0000-0003-2841-9738]]

[1] School of Computer Science, Technological University Dublin, Ireland
[2] Pavement Management Services Ltd. Athenry Co. Galway, Ireland
{ibrahim.syed, dympna.osullivan, susan.mckeever}@tudublin.ie

Abstract. Road maintenance and the early detection of road defects rely on routine pavement inspections. While advanced 3D laser profiling systems have the capability to automatically identify certain types of distress such as cracks and ruts, more complex pavement damage, including patches, often require manual identification. To address this limitation, this study proposes an automated patch detection system that employs object detection techniques. The results demonstrate the ability of object detection models to accurately identify patches in laser profiling images, indicating that the proposed approach has the capability to significantly enhance automation in visual inspection processes. This has the potential for significant cost reduction in inspections, improved safety conditions during checks, and acceleration of the current manual inspection processes.

Keywords: Road surface, visual inspection, Object detection, deep learning.

1 Introduction

Transportation and road infrastructure departments routinely conduct inspections on pavements to evaluate their surface conditions. Pavement conditions are impacted by traffic, weather, and sunlight, resulting in rutting, cracking, and ravelling that can ultimately lead to the disintegration of the surface layer. Such deterioration may be limited to the surface or indicative of more serious underlying structural issues related to the pavement. The maintenance of pavement surfaces necessitates a substantial number of resources and capital to ensure that optimal maintenance treatment is performed at the appropriate time. These inspections support informed decisions regarding pavement maintenance planning, including cost considerations [1]. In addition, the government can optimize allocation of limited resources for maintenance and consider long-term investment strategies [2]. Therefore, it is highly desirable to conduct pavement inspections in the most cost-effective manner.

Pavement assessment tasks involve a range of activities, from identifying pavement distresses to complete visual inspection of the pavement surface. To minimize maintenance costs, these tasks must align with the objectives of the schemes and the available budget. According to the survey, the UK, councils allocate 75% of funds for

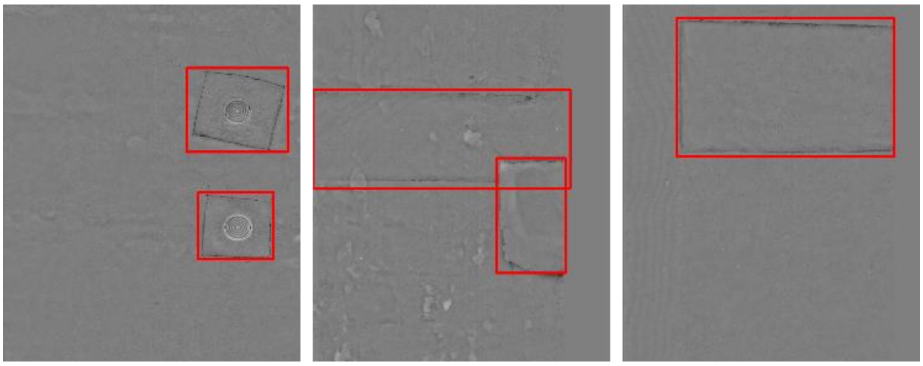


Fig. 1: Examples of laser profiling images acquired from asphalt pavement surface. **Note:** For the interpretation, patches are highlighted in the red boxes.

the maintenance of local road conditions and 25% for construction [3]. According to another survey, in 2008, around $182 billion was spent on capital improvements and maintenance of federal highways in the US [4].

1.1 Visual Inspection

There are two primary methods for visually inspecting pavements: manual and automatic. Visual inspection techniques typically involve three main steps: data acquisition, distress identification, and distress assessment. While the first step of data acquisition is largely automated using specialized vehicles, distress identification and assessment are typically performed manually. Manual inspection involves the assessment of pavement surface conditions by trained pavement engineers or certified inspectors, either through on-site surveys or by analyzing recorded videos/images. However, the manual visual inspection process is time-consuming, subjective, prone to errors, and potentially hazardous for inspectors.

To address these limitations, there is a growing interest in automating the entire visual inspection process. This can be achieved using advanced technologies such as machine learning, computer vision, and remote sensing [5].

1.2 Related Work

Several researchers have proposed the use of machine learning and computer vision-based methods to automate the distress identification step of visual pavement inspection. Implementing these techniques has the potential to significantly enhance the accuracy and efficiency of pavement inspections, while reducing the time and cost involved in manual inspections.

Various methods based on deep learning, particularly convolutional neural networks, have been suggested for the automatic identification of pavement distress. These methods can be broadly categorized into three main groups based on their approach to the problem of detecting pavement distress: classification, localization, and segmentation. For instance, Gupta et al. [6] present a pothole detection system using thermal images. The authors employed two object detection models i.e., the ResNet34-single shot multibox detector and the ResNet50-RetinaNet model. The ResNet34-single shot multibox detector yielded an average precision of 74.53%, while the ResNet50-RetinaNet model achieved a precision of 91.15%. Zou et al. [7] proposed a DeepCrack segmentation network that utilizes the encoder-decoder architecture to accurately detect pavement cracks. Through experimentation, the authors demonstrated that the proposed DeepCrack model achieves an F-measure of over 0.87, highlighting its efficacy in identifying and segmenting pavement image pixels into cracks and background.

Despite extensive research on automatic pavement distress detection, most studies have focused on potholes and cracks, with less attention given to the detection of pavement patches. The detection and localization of patches on pavement surfaces is an essential task for pavement management and maintenance. Patches play a dual role in the pavement maintenance process, offering both temporary and long-term pavement fix solutions and can be made of similar or different materials than the surrounding pavement. They serve various purposes, ranging from covering large areas like utility patching to addressing specific single-pothole distresses. Given the diverse nature of patching, manual intervention by engineers is necessary. They manually label and draw bounding boxes around each patch to aid in the detection process. The manual process is not only time consuming but also cost-intensive. In addition, modern 3D laser profiling technology, such as the LCMS (Laser Crack Measurement System) [8], and Pave3D [9] is now broadly utilized in the pavement inspection process. These systems are equipped with 3D laser profilers that capture range and intensity images of the pavement surface and come with a built-in data processing tool that processes the acquired data and automatically identifies various types of road cracks [10]. Furthermore, 3D laser profiling technology has become a standard technique in pavement inspection companies, where these systems are employed.

Advanced pavement inspection systems like Laser Crack Measurement System (LCMS) have made significant progress in automatically detecting various types of defects such as cracks [8], rutting [11], raveling [12], which can help in the assessment of pavement condition and the planning of maintenance and rehabilitation activities. However, automatic detection of pavement patches presents a significant challenge due to the similarity in color between intact pavement and patched surfaces. As a result, the detection process often involves manual involvement, with engineers manually labeling or drawing bounding boxes around each patch. Therefore, this paper addresses the problem of automatic patch detection on laser profiling images.

2 Methodology

This paper proposes a methodology for the detection and localization of patches using laser profiling images (see Figure 1). Our approach leverages state-of-the-art object detection techniques based on deep learning algorithms to achieve the detection of patch locations on pavement surfaces. This study aims to address the following research

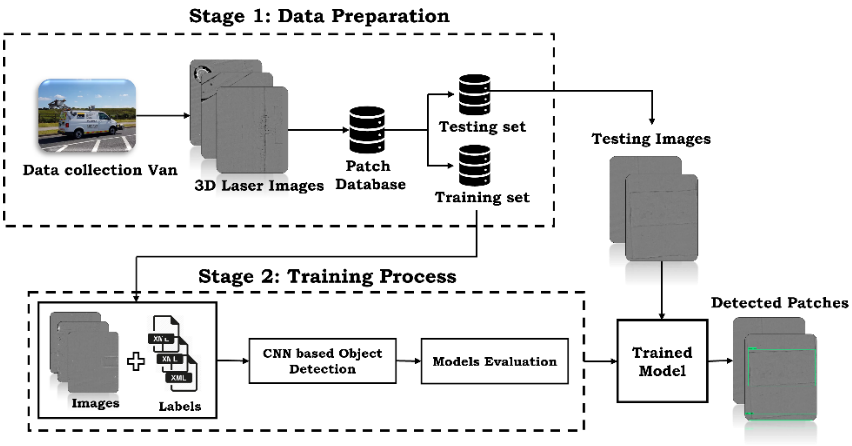


Fig. 2: Proposed pavement patch detection framework.

question: To what extent can an object detection model accurately detect and localize patches on the intact pavement? Figure 2 illustrates the proposed framework of automatic pavement patch detection system.

To address the above research question, we propose a comprehensive framework for an automatic pavement patch detection system (refer to Figure 2). Our methodology involves a multi-stage process, which includes pre-processing of the 3D laser profiling images, training and fine-tuning of the object detection model, and post-processing of the detected patches. The pre-processing step involved collecting a dataset of 3D laser profiling images that captured the pavement surface. From this dataset, we selected positive images only (i.e., images that contain one or more patches). These positive images were then manually labeled by an expert by drawing bounding boxes around the patches, indicating their locations in the images. We utilize a state-of-the-art object detection model based on deep learning algorithms, which have been proven to outperform traditional machine learning and images processing techniques [13]. Once the object detection model identified and localized the patches in the images, in the post-processing step the detected coordinates of the bounding boxes were used to calculate the area of the patched surface on the intact pavement surface. These calculated areas then allowed us to quantify the extent of patching on the normal pavement surface. In addition, given that the patch detection models trained on data obtained from a single country and specific capture settings, Therefore, in order to evaluate the generalizability of the model to road surfaces in different regions with different capture settings, we conducted tests on images acquired from the road network in another country (Ireland).

2.1 Data Collection & Preparation

This study employs asphalt pavement images acquired through 3D laser profiling system LCMS (Laser Crack Measurement System) mounted on the back of a data collection van. This system employs high-speed and high-resolution transverse profiling to capture images of pavement surface. The LCMS takes a transverse profile every 5mm. That means every 5mm, there are 4160 laser readings taken across the width of the pavement at 1mm increments. A transverse profile every 5mm gives a good profile of the road surface and allows a survey speed of around 90km/h. Each line in the depth image represents 5mm of road length, 200 lines represents one meter. The rate of transverse profile collection can be adjusted but 5mm is a good compromise of speed, safety, and data collected. The laser profiling system generates range image output, with sample images shown in Figure 1. Furthermore, each image has been manually labelled by an engineer at PMS (Pavement Management Services, Ireland) by drawing bounding boxes around the patches in the image.

In this study, 80% of the data was utilized to train the model, while the remaining 20% was reserved for evaluating the model's performance. The dataset details are described in Table 1.

Table 1: Details of the complete training and testing sets.

# Of images	Training set	Testing set
1874	1500	374

2.2 Network Architecture

This study employed two distinct network architectures to obtain comparative results utilizing the specified dataset. Specifically, the state-of-the-art object detection models Faster RCNN [14] and DETR [15] were selected for training and testing using transfer learning technique [16]. The transfer learning technique known as "fine-tuning" was employed to adapt a pre-trained Faster RCNN model to our specific object detection task. Specifically, we utilized a Faster RCNN model pre-trained on a large-scale dataset, such as ImageNet, which had learned general visual features. We then initialized our model with these pre-trained weights and further fine-tuned it on our custom dataset with annotations for the target objects. By fine-tuning, we allowed the model to adapt its learned features to our domain-specific data, enabling it to perform object detection more effectively and efficiently for our specific task. Furthermore, the selection of Faster RCNN and DETR models was motivated by their prior usage in the domain of automatic pavement inspection, including applications such as road marking detection [17], pothole detection [18], and other surface distress identification [19].

2.3 Evaluation Protocol

To assess the effectiveness of the trained model in detecting and localizing patches, its performance was evaluated using the mean Average Precision (mAP) score, specifically for the patch category. Average precision is determined by calculating the area

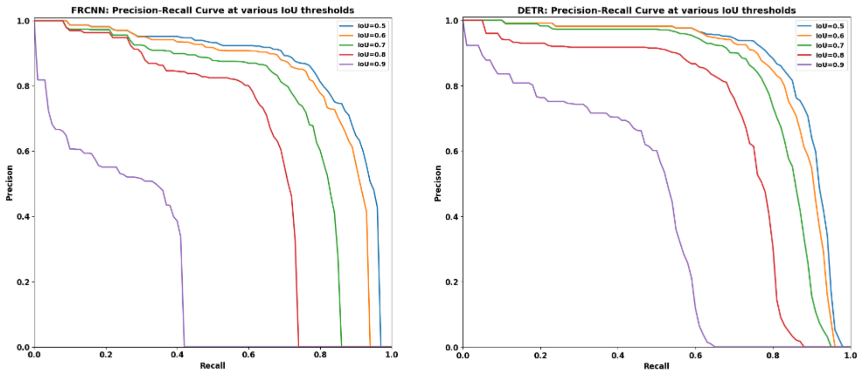


Fig. 3: Precision-recall curve of both models obtained at highest and lowest IoU thresholds.

under the Precision-Recall curve for each category, providing a comprehensive understanding of the model's performance across the range of precision and recall values. To achieve the mean average precision score, an IoU threshold of 0.5 was established. This threshold is a widely accepted evaluation metric employed by the PASCAL VOC object detection competition [20]. Furthermore, according to the domain expert for the task of patch detection, a higher IoU threshold is not required, as the exact placement of the patch relative to the predicted area only needs to be enough to say that a patch exists in the intact pavement area [21]. The mean average precision, precision and recall can be calculated as follow:

$$mAP = \frac{1}{N_{classes}} \sum_i AP_i \quad (1)$$

Where i is a type of distress and $N_{classes}$ is the total number of distress types, which is 1 in our case.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

Where True positives + False positives is the total number of detections generated from the model.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (3)$$

Where True positives + False negatives is the total number of ground truth boxes.

2.4 Experimental Results

The detection performance of the two models is presented in Table 2, while Figure 3 shows the precision-recall curve obtained at the highest and lowest IoU thresholds. The

DETR model outperforms the Faster RCNN model, with a 2% increase in mean average precision.

Table 2: Performance of both models on the test set.

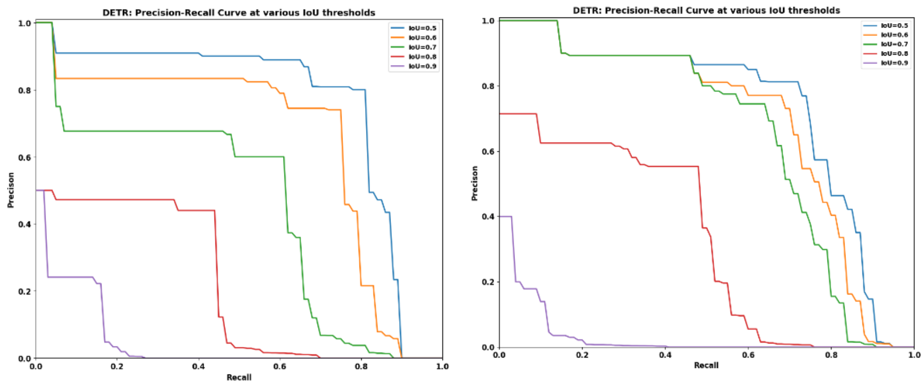
Model	**Backbone**	**$mAP_{0.5}$**
Faster RCNN	ResNet50	0.86
DETR	ResNet50	0.88

The analysis of the precision-recall graphs (see Figure-3) revealed that a higher IoU threshold led to a decrease in recall rate. This outcome is because the model only identifies patches as true positives if their overlap is greater than or equal to 0.9. Alternatively, using an IoU of 0.5 would result in an increase in recall rate as shown in Figure 3 Furthermore, from the precision and recall graphs (Figure 3) we can easily find the best possible precision and recall value. For instance, the blue line on the graph indicates the results obtained at the 0.5 IoU threshold. This line allows us to easily determine where the ideal precision and recall values can be located. For instance, at a recall rate of 0.8, the model's precision will be approximately 0.92. Using this analysis, we can identify the best precision and recall values using different IoU thresholds. It is important to note that if we maintain a tight IoU threshold, the model's precision and recall will gradually decrease. For example, the purple line represents the results achieved at a 0.9 IoU. At a recall rate of 0.4, the precision of the model will be approximately 0.72.

Overall, our results demonstrate the effectiveness of both models in pavement patch detection, in the context of the evaluation performed. However, it is important to note that the choice of evaluation metric and IoU threshold can have a significant impact on the results, and different scenarios may require different models or parameter settings. Additionally, the suitability of deploying such systems in real-world applications largely depends on the task requirements and priorities of the task owner. For instance, if the cost of missing a patch is high, a lower false negative rate may be necessary, even at the cost of an increased false positive rate. Conversely, a false positive rate may be tolerable, provided that the false negative rate remains sufficiently low. In summary, the ideal configuration of object detection models in real-world applications should consider the specific trade-offs between false positives and false negatives that align with the task's objectives and constraints. In addition, as per the insights of domain experts at PMS, it is imperative to consider this trade-off within the specific scope of the patch detection task at hand. To achieve higher recall, the acceptance of false positives may be deemed acceptable.

2.5 Patch Detection on Different Pavement Conditions

The aim of this analysis is to assess the performance of trained models in the context of altered conditions, including variations in pavement surface and image capture



(A) 10mm Interval **(B) 5mm Interval**

Fig. 4: Precision-recall curves achieved by using best performing model, with two different intervals of images: (A) shows the results achieved using 10mm interval images, while (B) shows the results achieved using 5mm interval.

settings. Our aim is to evaluate the effectiveness of the Faster RCNN and DETR models, which were originally trained on images obtained from the USA road network utilizing a laser profiling system set to acquire images at 5mm intervals. To achieve this goal, we carried out an experiment on the Ireland road network utilizing the same models, but with the laser profiling system adjusted to capture images at 10mm intervals (see Figure 5). Table 3 demonstrates the results achieved by both models on 10mm interval images.

Our experimental investigation involved the analysis 41 images captured at 10mm intervals on the Irish road network. The findings indicate that the model performance decreased by 13~15% compared to the results achieved on test set (see Table 2) due to shift in image resolution compared to the images used for model training. Additionally, the shape and composition of the pavement surface may have contributed to a decrease in overall results across both interval images. Table 3 demonstrates the results of further analysis, indicating that there is a slight difference between 10mm interval images results and 5mm interval image segments results. The reason for the variance between the 10mm images and 5mm image segments is the difference in the number of images. With only 41 images at 10mm intervals, breaking them down into 5mm segments results in 82 images, which increases the number of ground truth boxes. This factor may contribute to the observed differences in model performance between the two image types. The overall findings suggest that changes in image quality, patch shape, and pavement surface have a minor impact on model performance, and that additional training or calibration may be necessary for the models to operate effectively in environments with distinct imaging conditions. Moreover, Figure 4 depicts a precision and recall curve that can aid in identifying the ideal IoU and confidence threshold values for obtaining the desired results in real-world testing scenario. Figure 6 shows in some cases where the model fails to detect the true patch within an image interval of 10mm, but subsequently identifies it in the corresponding 5mm interval images.

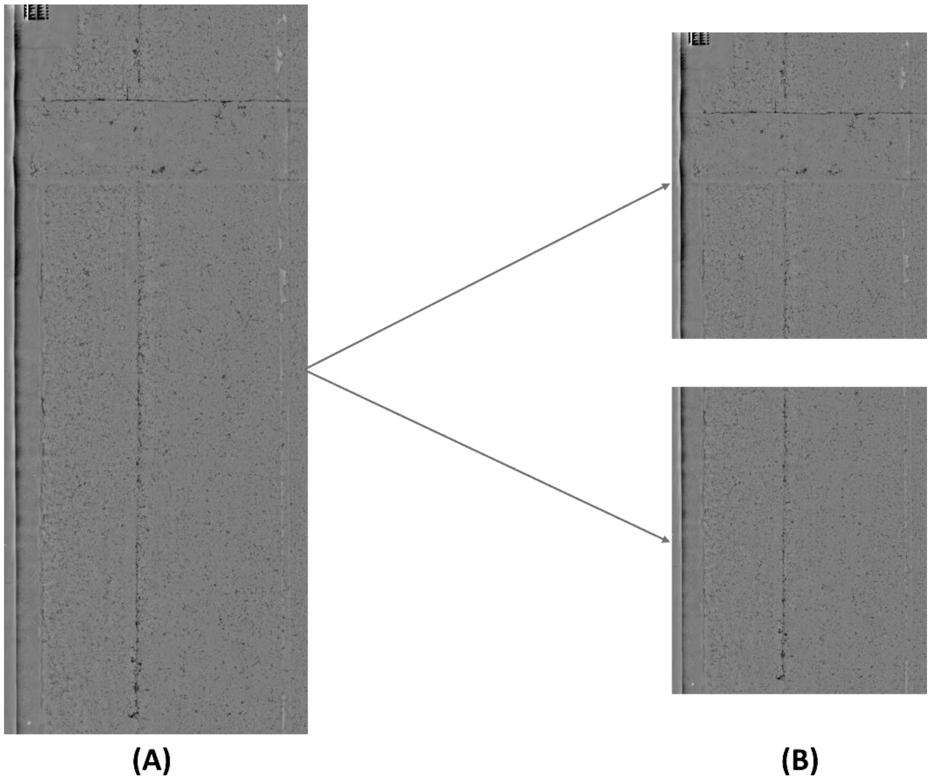


Fig. 5: 10mm image (A) and its corresponding 5mm segments (B).

Table 3: Comparative analysis of both models on 10mm & 5mm interval images.

Model	Interval	# Of images	# Of Patches	mAP
FRCNN	10mm	41	49	0.72
DETR				0.75
FRCNN	5mm	82	54	0.7
DETR				0.73

Figure 7 illustrates the visual results predicted by both models. This can be attributed to the fact that the model was trained on 5mm images, which makes it easier to detect patches in such images. Secondly, the shift in image resolution also contributes to the model's difficulty in detecting patches on 10mm interval images.

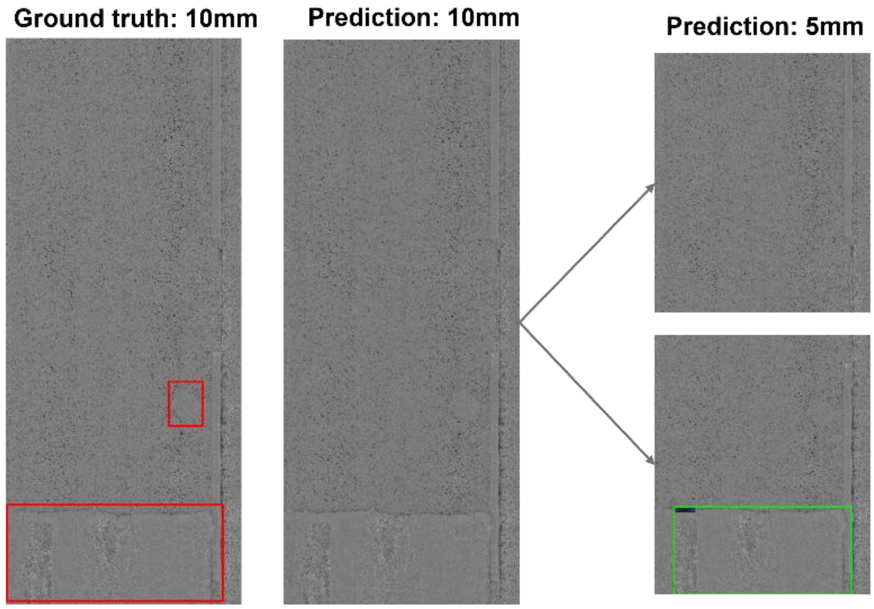


Fig. 6: Visual results of 10mm interval image: Ground truth, prediction on 10mm image and its corresponding 5mm interval images.

3 Conclusion

This paper addresses automated patch detection, as a step towards an automated system for detecting asphalt pavement patches through the utilization of laser profiling images. Two state-of-the-art object detection models, Faster RCNN and DETR, were trained using a specified dataset. The results indicate that both models achieved higher patch detection accuracy, with a mean average precision of 0.86 and 0.88 for the Faster RCNN and DETR models, respectively. Compared to the Faster RCNN, the DETR model showed slightly improved performance. Additionally, we assessed the model's generalization capabilities under different road surface and image capture conditions and discovered that the model's performance was slightly affected by the variability in conditions, such as change in the pavement surface, and image capture settings. Furthermore, given consideration to the trade-off between precision and recall is crucial in any object detection task. Different applications may prioritize precision (minimizing false positives) or recall (minimizing false negatives) differently, depending on the specific use case and requirements. The trade-off between precision and recall, as highlighted by industry domain experts at PMS, is crucial to consider within the specific context of patch detection work. For example, in patch detection application scenarios, it is deemed acceptable to tolerate false positives in order to achieve a higher recall rate. This emphasizes the importance of aligning the model's performance with the specific requirements and priorities of the task at hand.

Future work will extend the study to calculate the area of each identified patch and analyse the severity of patched surface.

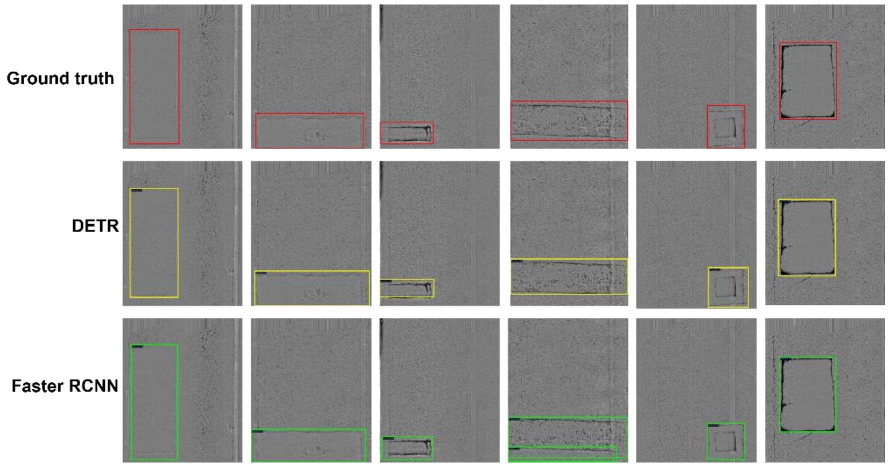


Fig. 7: Visual results of 5mm interval images: Ground truth, DETR prediction and Faster RCNN prediction.

Acknowledgment

This work was funded by Science Foundation Ireland through the SFI Centre for Research Training in Machine Learning (18/CRT/6183).

References

1. C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," *Adv. Eng. Informatics*, vol. 25, no. 3, pp. 507–515, Aug. 2011, doi: 10.1016/j.aei.2011.01.002.
2. C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, and P. Fieguth, "A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure," *Adv. Eng. Informatics*, vol. 29, no. 2, pp. 196–210, 2015.
3. S. C. Radopoulou and I. Brilakis, "Patch detection for pavement assessment," *Autom. Constr.*, vol. 53, pp. 95–104, 2015.
4. E. Schnebele, B. F. Tanyu, G. Cervone, and N. Waters, "Review of remote sensing methodologies for pavement management and assessment," *Eur. Transp. Res. Rev.*, vol. 7, no. 2, pp. 1–19, 2015.
5. N. Sholevar, A. Golroo, and S. R. Esfahani, "Machine learning techniques for pavement condition evaluation," *Autom. Constr.*, vol. 136, p. 104190, 2022.
6. S. Gupta, P. Sharma, D. Sharma, V. Gupta, and N. Sambyal, "Detection and localization of potholes in thermal images using deep neural networks," *Multimed. Tools Appl.*, vol. 79, no. 35, pp. 26265–26284, 2020.

7. Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, “Deepcrack: Learning hierarchical convolutional features for crack detection,” *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1498–1512, 2018.
8. Laurent, J., Fox-Ivey, R., & Petitclerc, B. (2020). High resolution multi-lane road surface mapping using 3D laser profilers for 3D paving and milling projects. In Proc., 7th Eurasphalt and Eurobitume Congress. Brussels, Belgium: European Asphalt Pavement Association.
9. W. Luo, K. C. P. Wang, L. Li, Q. J. Li, and M. Moravec, “Surface drainage evaluation for rigid pavements using an inertial measurement unit and 1-mm three-dimensional texture data,” *Transp. Res. Rec.*, vol. 2457, no. 1, pp. 121–128, 2014.
10. Laurent, J., Hébert, J. F., Lefebvre, D., & Savard, Y. (2012). Using 3D laser profiling sensors for the automated measurement of road surface conditions. In 7th RILEM International Conference on Cracking in Pavements: Mechanisms, Modeling, Testing, Detection and Prevention Case Histories (pp. 157-167). Springer Netherlands.
11. W. Luo, Y. Qin, D. Zhang, and L. Li, “Measurement of pavement rutting trajectories on two-lane highway using the 3D line scanning laser system,” *Int. J. Pavement Eng.*, pp. 1–16, 2022.
12. S. Mathavan, M. M. Rahman, M. Stonecliffe-Janes, and K. Kamal, “Pavement raveling detection and measurement from synchronized intensity and range images,” *Transportation Research Record*, vol. 2457. National Research Council, pp. 3–11, 2014, doi: 10.3141/2457-01.
13. W. S. Qureshi *et al.*, “An Exploration of Recent Intelligent Image Analysis Techniques for Visual Pavement Surface Condition Assessment,” *Sensors*, vol. 22, no. 22, p. 9019, 2022.
14. S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016.
15. N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-end object detection with transformers,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16*, 2020, pp. 213–229.
16. K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J. Big data*, vol. 3, no. 1, pp. 1–40, 2016.
17. H. Alzraiee, A. Leal Ruiz, and R. Sprotte, “Detecting of pavement marking defects using faster R-CNN,” *J. Perform. Constr. Facil.*, vol. 35, no. 4, p. 4021035, 2021.
18. S. I. Hassan, D. O’Sullivan, and S. Mckeever, “Pothole Detection under Diverse Conditions using Object Detection Models.,” *IMPROVE*, vol. 1, pp. 128–136, 2021.
19. J. Wang, G. Xu, F. Yan, J. Wang, and Z. Wang, “Defect transformer: An efficient hybrid transformer architecture for surface defect detection,” *Measurement*, p. 112614, 2023.
20. M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes (voc) challenge,” *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, 2010.
21. S. I. Hassan, D. O’Sullivan, S. McKeever, D. Power, R. McGowan, and K. Feighan, “Detecting Patches on Road Pavement Images Acquired with 3D Laser Sensors using Object Detection and Deep Learning.,” in *VISIGRAPP (5: VISAPP)*, 2022, pp. 413–420.